Week 10 Markdown

Dylan Denner

3/30/2021

## R Markdown

# Simple Linear Regression Models  
  
POVERTY\_lm <- lm(HS\_PLUS\_percentage ~ poverty\_percentage, data = mod\_model\_data)  
DISCIPLINE\_lm <- lm(HS\_PLUS\_percentage ~ mean\_total\_students\_discipline, data = mod\_model\_data)  
CHRONIC\_lm <- lm(HS\_PLUS\_percentage ~ mean\_chronic\_absenteeism, data = mod\_model\_data)  
ATTENDANCE\_lm <- lm(HS\_PLUS\_percentage ~ mean\_attendance, data = mod\_model\_data)  
ENROLLMENT\_lm <- lm(HS\_PLUS\_percentage ~ mean\_enrollment, data = mod\_model\_data)

Shapiro-Wilk Normality Test for Poverty %

POVERTY\_sresid <- studres(POVERTY\_lm)  
shapiro.test(POVERTY\_sresid)

##   
## Shapiro-Wilk normality test  
##   
## data: POVERTY\_sresid  
## W = 0.95023, p-value = 3.341e-13

Shapiro-Wilk Normality Test for Mean Student Discipline

DISCIPLINE\_sresid <- studres(DISCIPLINE\_lm)  
shapiro.test(DISCIPLINE\_sresid)

##   
## Shapiro-Wilk normality test  
##   
## data: DISCIPLINE\_sresid  
## W = 0.95965, p-value = 1.215e-11

Shapiro-Wilk Normality Test for Mean Chronic Absenteeism %

CHRONIC\_sresid <- studres(CHRONIC\_lm)  
shapiro.test(CHRONIC\_sresid)

##   
## Shapiro-Wilk normality test  
##   
## data: CHRONIC\_sresid  
## W = 0.94531, p-value = 6.133e-14

Shapiro-Wilk Normality Test for Mean Attendance %

ATTENDANCE\_sresid <- studres(ATTENDANCE\_lm)  
shapiro.test(ATTENDANCE\_sresid)

##   
## Shapiro-Wilk normality test  
##   
## data: ATTENDANCE\_sresid  
## W = 0.94197, p-value = 2.056e-14

Shapiro-Wilk Normality Test for Mean Enrollment

ENROLLMENT\_sresid <- studres(ENROLLMENT\_lm)  
shapiro.test(ENROLLMENT\_sresid)

##   
## Shapiro-Wilk normality test  
##   
## data: ENROLLMENT\_sresid  
## W = 0.96928, p-value = 8.867e-10

Score Test for Non-Constant Error Variance: Poverty %

#heterodastic  
ncvTest(POVERTY\_lm)

## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 26.53847, Df = 1, p = 2.5834e-07

Score Test for Non-Constant Error Variance: Mean Student Discipline

#homoscedastic  
ncvTest(DISCIPLINE\_lm)

## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 0.08208011, Df = 1, p = 0.7745

Score Test for Non-Constant Error Variance: Mean Chronic Absenteeism %

#heteroscedastic  
ncvTest(CHRONIC\_lm)

## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 6.226651, Df = 1, p = 0.012584

Score Test for Non-Constant Error Variance: Mean Attendance %

#homoscedastic  
ncvTest(ATTENDANCE\_lm)

## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 0.00171754, Df = 1, p = 0.96694

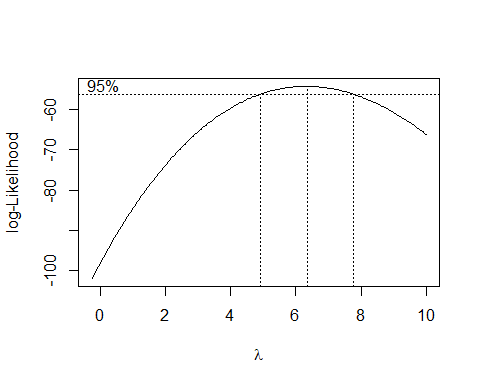
Score Test for Non-Constant Error Variance: Mean Enrollment

#heteroscedastic  
ncvTest(ENROLLMENT\_lm)

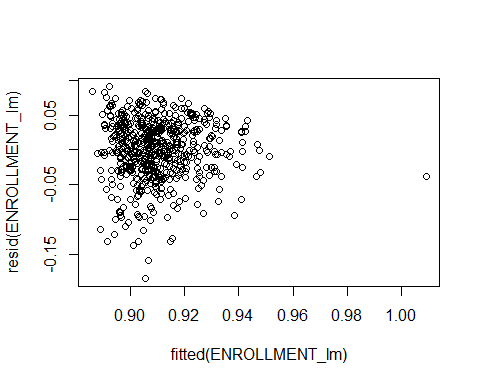
## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 11.81214, Df = 1, p = 0.00058846

BOXCOX Transformations

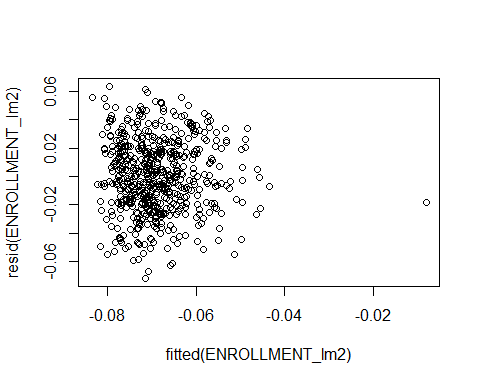
library(MASS)  
boxcox(ENROLLMENT\_lm, lambda = seq(-0.25, 10, by = 0.05), plotit = TRUE)



plot1 <- plot(fitted(ENROLLMENT\_lm), resid(ENROLLMENT\_lm))



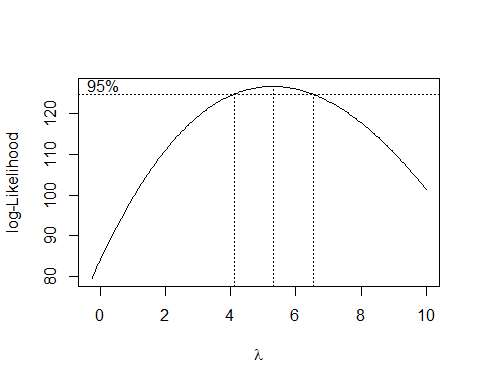
ENROLLMENT\_lm2 <- lm((((HS\_PLUS\_percentage ^ 6) - 1) / 6) ~ mean\_enrollment, data = mod\_model\_data)  
plot2 <- plot(fitted(ENROLLMENT\_lm2), resid(ENROLLMENT\_lm2))



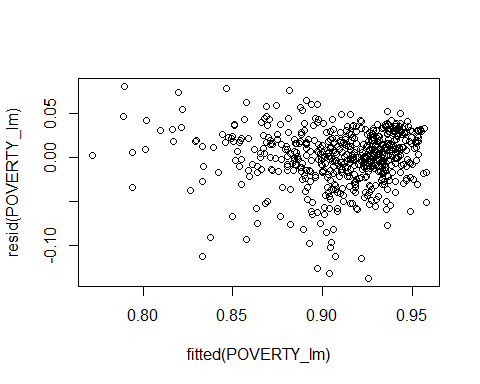
ENROLLMENT\_sresid2 <- studres(ENROLLMENT\_lm2)  
shapiro.test(ENROLLMENT\_sresid2)

##   
## Shapiro-Wilk normality test  
##   
## data: ENROLLMENT\_sresid2  
## W = 0.99584, p-value = 0.1196

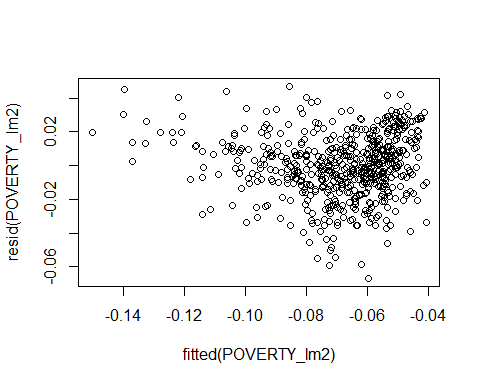
#Better P-value for normality, but still significantly Different from Normal  
POVERTY\_lm <- lm(HS\_PLUS\_percentage ~ poverty\_percentage, data = mod\_model\_data)  
boxcox(POVERTY\_lm, lambda = seq(-0.25, 10, by = 0.05), plotit = TRUE)



plot1 <- plot(fitted(POVERTY\_lm), resid(POVERTY\_lm))



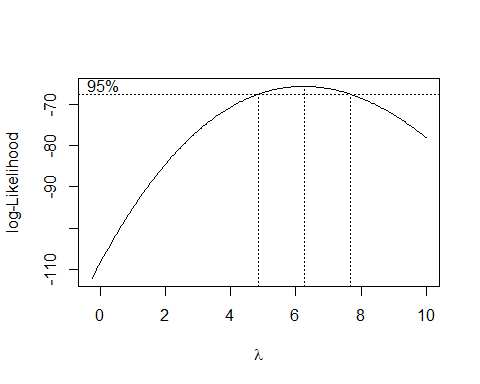
POVERTY\_lm2 <- lm((((HS\_PLUS\_percentage ^ 6) - 1) / 6) ~ poverty\_percentage, data = mod\_model\_data)  
plot2 <- plot(fitted(POVERTY\_lm2), resid(POVERTY\_lm2))



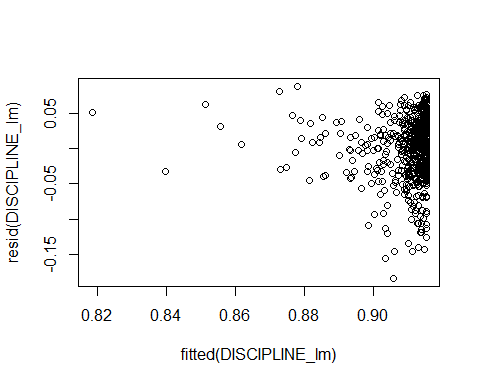
POVERTY\_sresid2 <- studres(POVERTY\_lm2)  
shapiro.test(POVERTY\_sresid2)

##   
## Shapiro-Wilk normality test  
##   
## data: POVERTY\_sresid2  
## W = 0.99205, p-value = 0.003028

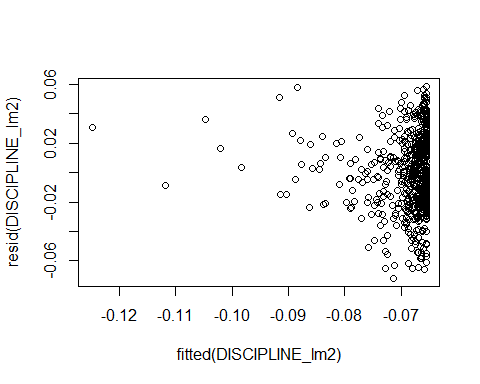
#Better P-value for normality, but still significantly Different from Normal  
DISCIPLINE\_lm <- lm(HS\_PLUS\_percentage ~ mean\_total\_students\_discipline, data = mod\_model\_data)  
boxcox(DISCIPLINE\_lm, lambda = seq(-0.25, 10, by = 0.05), plotit = TRUE)



plot1 <- plot(fitted(DISCIPLINE\_lm), resid(DISCIPLINE\_lm))



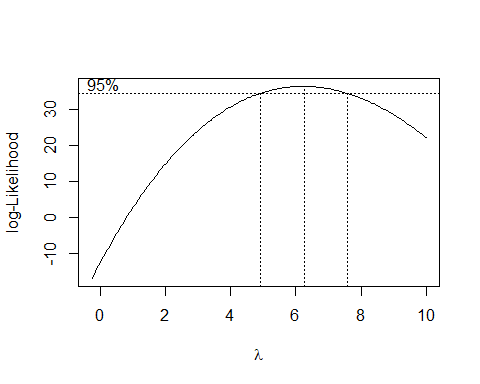
DISCIPLINE\_lm2 <- lm((((HS\_PLUS\_percentage ^ 6) - 1) / 6) ~ mean\_total\_students\_discipline, data = mod\_model\_data)  
plot2 <- plot(fitted(DISCIPLINE\_lm2), resid(DISCIPLINE\_lm2))



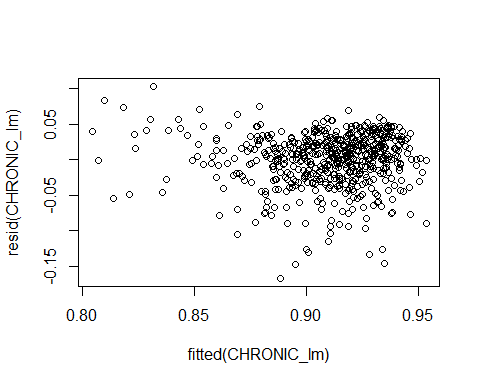
DISCIPLINE\_sresid2 <- studres(DISCIPLINE\_lm2)  
shapiro.test(DISCIPLINE\_sresid2)

##   
## Shapiro-Wilk normality test  
##   
## data: DISCIPLINE\_sresid2  
## W = 0.99339, p-value = 0.0108

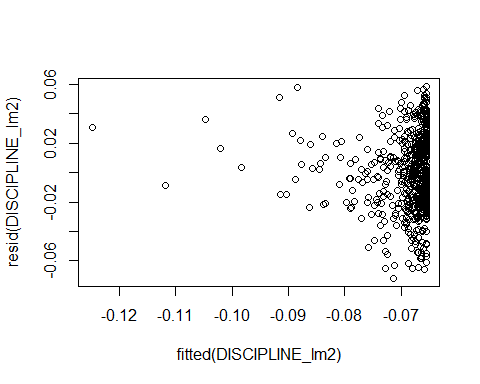
#Better P-value for normality, but still significantly Different from Normal  
CHRONIC\_lm <- lm(HS\_PLUS\_percentage ~ mean\_chronic\_absenteeism, data = mod\_model\_data)  
boxcox(CHRONIC\_lm, lambda = seq(-0.25, 10, by = 0.05), plotit = TRUE)



plot1 <- plot(fitted(CHRONIC\_lm), resid(CHRONIC\_lm))



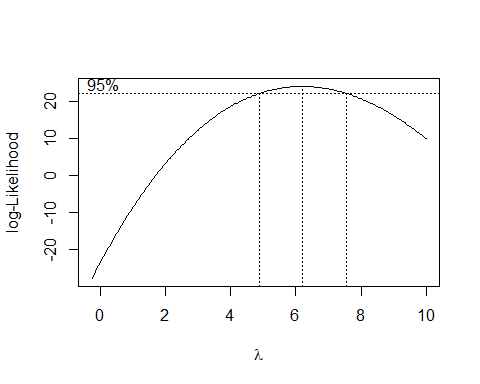
CHRONIC\_lm2 <- lm((((HS\_PLUS\_percentage ^ 6) - 1) / 6) ~ mean\_chronic\_absenteeism, data = mod\_model\_data)  
plot2 <- plot(fitted(DISCIPLINE\_lm2), resid(DISCIPLINE\_lm2))



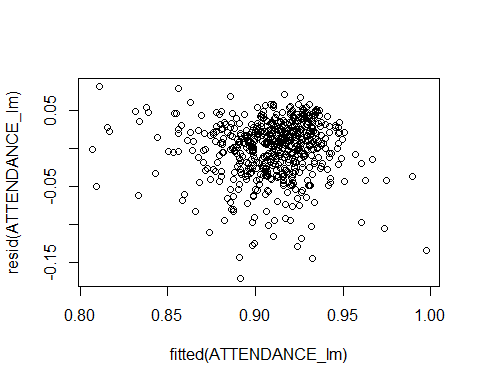
CHRONIC\_sresid2 <- studres(CHRONIC\_lm2)  
shapiro.test(DISCIPLINE\_sresid2)

##   
## Shapiro-Wilk normality test  
##   
## data: DISCIPLINE\_sresid2  
## W = 0.99339, p-value = 0.0108

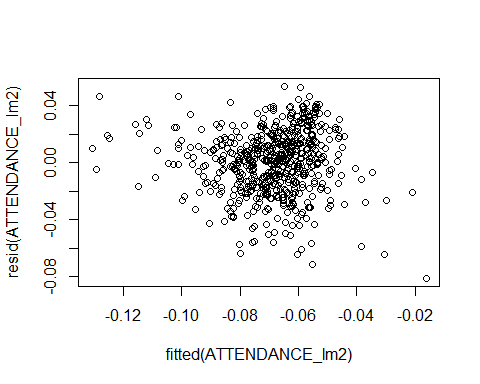
#Better P-value for normality, but still significantly Different from Normal  
ATTENDANCE\_lm <- lm(HS\_PLUS\_percentage ~ mean\_attendance, data = mod\_model\_data)  
boxcox(ATTENDANCE\_lm, lambda = seq(-0.25, 10, by = 0.05), plotit = TRUE)



plot1 <- plot(fitted(ATTENDANCE\_lm), resid(ATTENDANCE\_lm))



ATTENDANCE\_lm2 <- lm((((HS\_PLUS\_percentage ^ 6) - 1) / 6) ~ mean\_attendance, data = mod\_model\_data)  
plot2 <- plot(fitted(ATTENDANCE\_lm2), resid(ATTENDANCE\_lm2))



ATTENDANCE\_sresid2 <- studres(ATTENDANCE\_lm2)  
shapiro.test(ATTENDANCE\_sresid2)

##   
## Shapiro-Wilk normality test  
##   
## data: ATTENDANCE\_sresid2  
## W = 0.98915, p-value = 0.0002359